# Executive Summary: DICK’S Sporting Goods Scraping Feasibility

## Overview

**Target**: https://www.dickssportinggoods.com/  
**Assessment Date**: October 8, 2025  
**Difficulty Rating**: 9/10 (HARD)  
**Recommended Approach**: Browser Automation (Mandatory)

## Key Findings

### Protection Assessment

DICK’S Sporting Goods implements **enterprise-grade bot protection** with Akamai Bot Manager that completely blocks HTTP-based scraping attempts. Even with authentic browser headers extracted from live sessions, HTTP requests achieve a **0% success rate**, making browser automation the only viable approach.

### Product Catalog Scale

* **Total Products**: ~179,651 products across extensive categories
* **Product Depth**: Comprehensive sporting goods, apparel, footwear, and outdoor equipment
* **Data Richness**: Complete product specifications, pricing, inventory, and multimedia content
* **Business Value**: High-value competitive intelligence opportunity

### Technical Requirements

* **Browser Automation**: Mandatory - Playwright/Selenium with stealth configurations
* **Proxy Infrastructure**: Residential proxies required (datacenter proxies fail)
* **Success Rate**: 85-95% with proper browser automation and proxy rotation
* **Resource Intensity**: High CPU/memory requirements, 24/7 monitoring needed

## Business Impact Assessment

### Opportunity Value

* **Market Intelligence**: Access to pricing strategies across 179K+ products
* **Competitive Analysis**: Monitor product launches, promotions, and inventory patterns
* **Brand Performance**: Track major sporting goods brands and category trends
* **Revenue Potential**: Significant competitive advantage in sporting goods market

### Investment Requirements

* **Infrastructure**: Enterprise-grade cloud infrastructure with auto-scaling
* **Proxy Services**: Premium residential proxy pools ($3,000-5,000/month)
* **Maintenance**: Dedicated anti-bot specialists and 24/7 monitoring
* **Timeline**: 6-8 weeks for complete implementation and testing

## Risk Analysis

### Technical Risks

* **High Complexity**: Sophisticated multi-layered protection systems
* **Maintenance Overhead**: Continuous adaptation required as protections evolve
* **Success Variability**: Protection updates can temporarily impact operations
* **Resource Intensity**: Significant infrastructure and operational costs

### Operational Risks

* **Legal Considerations**: Must comply with terms of service and rate limiting
* **Detection Risk**: Advanced behavioral analysis and fingerprinting active
* **Scalability Challenges**: Limited by bot protection sophistication
* **Dependency Risk**: Success depends on advanced proxy and automation services

## Strategic Recommendations

### Immediate Actions

1. **Feasibility Confirmation**: Approve high-difficulty, high-investment project
2. **Resource Allocation**: Assign dedicated development team with anti-bot expertise
3. **Infrastructure Planning**: Design cloud-based auto-scaling architecture
4. **Proxy Procurement**: Establish relationships with premium residential proxy providers

### Implementation Strategy

1. **Phase 1**: Proof of concept with 5,000-10,000 high-value products
2. **Phase 2**: Scale to 50,000 products across major categories
3. **Phase 3**: Full catalog coverage with optimized monitoring and maintenance
4. **Monitoring**: Implement real-time success rate tracking and alerting

### Success Metrics

* **Data Completeness**: Target 95% of available product information
* **Success Rate**: Maintain 85%+ successful extraction rate
* **Coverage**: Complete catalog scan every 6-12 days
* **Performance**: 15,000-25,000 products processed daily

## Alternative Considerations

### Partnership Approach

Consider official data partnership discussions with DICK’S Sporting Goods for legitimate business use cases, which may provide more stable access than scraping.

### Selective Scraping

Focus on highest-value categories (Nike, Adidas, premium brands) to maximize ROI while minimizing technical complexity and resource requirements.

### API Investigation

Research potential undocumented APIs or mobile app endpoints that might offer more accessible data sources.

## Financial Projections

### Estimated Monthly Costs

* **Proxy Services**: $3,000-5,000
* **Infrastructure**: $2,000-4,000
* **Maintenance & Monitoring**: $5,000-8,000
* **Total Monthly**: $10,000-17,000

### Return on Investment

* **Data Value**: High competitive intelligence value
* **Market Advantage**: Significant pricing and trend insights
* **Break-even**: 6-12 months depending on business application
* **Long-term Value**: Sustained competitive advantage in sporting goods sector

## Decision Framework

### Proceed If:

* ✅ Budget supports $10K-17K monthly operational costs
* ✅ Team has advanced anti-bot and browser automation expertise
* ✅ Business case justifies high-complexity, high-value data extraction
* ✅ Organization can commit to ongoing maintenance and adaptation

### Consider Alternatives If:

* ❌ Budget constraints limit infrastructure investment
* ❌ Limited technical expertise in advanced bot evasion
* ❌ Preference for lower-risk, lower-maintenance solutions
* ❌ Uncertain about long-term data usage and ROI

## Conclusion

DICK’S Sporting Goods represents a **high-reward, high-complexity** scraping opportunity. The 179,651-product catalog offers exceptional competitive intelligence value, but success requires **significant investment** in browser automation infrastructure, residential proxies, and specialized expertise.

**Recommendation**: Proceed only with **full commitment** to enterprise-grade implementation, understanding that this is among the most challenging e-commerce scraping targets due to Akamai Bot Manager protection.

**Success Probability**: 85-90% with proper resources and expertise  
**Investment Level**: High ($120K-200K annually)  
**Strategic Value**: Very High for sporting goods market intelligence